No evidence of inbreeding depression in fast declining herds of migratory caribou

Marianne Gagnon, Glenn Yannic, Charles Perrier, Steeve Côté

To cite this version:

HAL Id: hal-02921343
https://hal.inrae.fr/hal-02921343
Submitted on 25 Aug 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License
No evidence of inbreeding depression in fast declining herds of migratory caribou

Marianne Gagnon | Glenn Yannic | Charles Perrier | Steeve D. Côté

Abstract
Identifying inbreeding depression early in small and declining populations is essential for management and conservation decisions. Correlations between heterozygosity and fitness (HFCs) provide a way to identify inbreeding depression without prior knowledge of kinship among individuals. In Northern Quebec and Labrador, the size of two herds of migratory caribou (Rivière-George, RG and Rivière-aux-Feuilles, RAF) has declined by one to two orders of magnitude in the last three decades. This raises the question of a possible increase in inbreeding depression originating from, and possibly contributing to, the demographic decline in those populations. Here, we tested for the association of genomic inbreeding indices (estimated with 22,073 SNPs) with body mass and survival in 400 caribou sampled in RG and RAF herds between 1996 and 2016. We found no association of individual heterozygosity or inbreeding coefficient with body mass or annual survival. Furthermore, those genomic inbreeding indices remained stable over the period monitored. These results suggest that the rapid and intense demographic decline of the herds did not cause inbreeding depression in those populations. Although we found no evidence for HFCs, if demographic decline continues, it is possible that such inbreeding depression would be triggered.

KEYWORDS
general effect hypothesis, heterozygosity–fitness correlation, inbreeding depression, migratory caribou, Rangifer tarandus, SNP

1 | INTRODUCTION

Inbreeding depression, the reduced fitness of inbred individuals (Charlesworth & Charlesworth, 1987), can contribute to genetic Allee effects, a positive correlation between population size and fitness (Luque et al., 2016; Wittmann, Stuis, & Metzler, 2018), and can hence increase the risk of extinction of small and/or declining populations. Because inbred individuals have an increased probability of expressing recessive deleterious mutations or homozygous genotypes at over-dominant genes (Charlesworth & Charlesworth, 1987), they usually have low survival (Cecchi, Giacalone, & Paci, 2016; Norén, Godøy, Dalén, Meijer, & Angerbjörn, 2016), reproduction (Norén et al., 2016) and/or poor body condition (Brommer, Kekkonen, & Wikström, 2015; Gholizadeh & Ghafouri-Kesbi, 2016). Lower values of life-history traits can in turn affect population growth (Gaillard, Festa-Bianchet, Yoccoz, Loison, & Toïgo, 2000). In a conservation context, it is thus essential to identify inbreeding depression early in a population and to understand how it affects individual fitness to avoid extinction vortexes (Gilpin & Soule, 1986; Tanaka, 1997, 1998, 2000).

Inbreeding depression is traditionally studied with the help of exhaustive pedigrees, requiring long-term monitoring of populations.
and extensive knowledge of many kinship relations among individuals (Cecchi et al., 2016; Norén et al., 2016; Silió, Barragán, Fernández, García-Casco, & Rodríguez, 2016). Although this approach directly assesses inbreeding, it is costly and often unrealistic to produce pedigrees for natural populations (but see Charmantier, Garant, & Kruuk, 2014). Because inbreeding increases genome-wide homozygosity (Wright, 1977), it is frequently assumed that individual heterozygosity can be used as a proxy for inbreeding level (Balloux, Amos, & Coulson, 2004; Miller et al., 2014). Heterozygosity–fitness correlations (HFCs) have thus become an appealing tool in conservation and evolutionary biology to indirectly detect inbreeding depression, without prior knowledge of kinship among individuals (Chapman, Nakagawa, Coltman, Slate, & Sheldon, 2009; Coltman & Slate, 2003; Hoffman et al., 2014). The rationale behind this approach is that a positive correlation between heterozygosity and fitness should appear when inbreeding depression affects a portion of the individuals in a population. In such cases, inbred individuals will have lower fitness as direct results of decreasing heterozygosity, whereas noninbred individuals will present higher heterozygosity and then higher fitness (general effect hypothesis; David, Delay, Berthou, & Jarne, 1995). Yet, recent studies have shown that genomic measures of relatedness better approximate pedigree-based inbreeding in comparison with multi-locus heterozygosity (Bérénos, Béhérous, Ellis, Pilkinson, & Pemberton, 2014; Huisman, Kruuk, Ellis, Clutton-Brock, & Pemberton, 2016; Kardos, Taylor, Ellegren, Luikart, & Allendorf, 2016). For instance, inbreeding coefficients based on measures of genomic relatedness such as $F_{\text{grm}}$ (Huisman et al., 2016; $F_{\text{HL}}$ in Yang, Lee, Goddard, & Visscher, 2011) that gives more weight to homozygosity of rare alleles proved to correlate well with fitness traits such as breeding success or offspring survival, even more so than pedigree-based inbreeding coefficients (Huisman et al., 2016).

Studies that have investigated the relationship between individual heterozygosity, as a proxy for inbreeding, and fitness usually found contrasted results (Chapman et al., 2009; Szulkin, Bierne, & David, 2010). In some cases, it has been shown that heterozygosity was positively associated with fitness- or performance-related traits (Brambilla, Biebach, Bassano, Bogliani, & von Hardenberg, 2015; Da Silva et al., 2009; Herdegen, Nadachowska-Brzyska, Konowalik, Babik, & Radwan, 2013; Hoffman et al., 2014) and those results were attributed to the general effect hypothesis (David et al., 1995). The general effect hypothesis posits that inbreeding is negatively correlated with (a) fitness and (b) genetic diversity (Slate et al., 2004; Szulkin et al., 2010). If the former assumption is broadly accepted (Allendorf, Aikten, & Luikart, 2013; Brommer et al., 2015; Charlesworth & Charlesworth, 1987; Norén et al., 2016), the later is still debated (Baloux et al., 2004; DeWoody & DeWoody, 2005; Miller et al., 2014). Indeed, it is argued that marker-based estimates of genetic diversity should reflect genome-wide diversity and that it requires at the very least a few hundreds genomic markers to properly assess global diversity (Balloux et al., 2004; DeWoody & DeWoody, 2005; Miller et al., 2014). Yet, HFC studies were until recently only based on a few microsatellite markers, usually between 5 and 15 loci (Coltman & Slate, 2003), as a surrogate of genome-wide diversity (Da Silva, Luikart, Yoccoz, Cohas, & Allainé, 2006; Herdegen et al., 2013; Velando, Barros, & Moran, 2015). Instead of demonstrating the effect of genome-wide diversity on fitness, those HFCs may result from the direct effect of diversity at functional loci or from local effect of fitness-related loci linked to the markers used to estimate diversity (direct or local effect hypothesis; David et al., 1995; David, 1998). Besides, some studies on HFCs using microsatellites showed that the diversity of only a few markers was associated with performance-related traits, providing support to the local effect hypothesis rather than the general effect hypothesis (García-Navas, Cáliz-Campal, Ferrer, Sanz, & Ortego, 2014; Guinand et al., 2013; Lieutenant-Gosselin & Bernatchez, 2006). The advent of next-generation sequencing techniques may facilitate the production of detailed studies on HFCs by giving access to many thousands single nucleotide polymorphisms (SNPs) at a relatively low cost, allowing accurate measures of genomic relatedness and heterozygosity (Hoffman et al., 2014; Huisman et al., 2016; Kardos et al., 2016; Miller et al., 2014; Mitchell, Vitikainen, Wells, Cant, & Nichols, 2017).

In Northern Quebec and Labrador (Canada), Riviére-George (RG) and Riviére-aux-Feuilles (RAF) herds of migratory caribou (Rangifer tarandus; called reindeer in Eurasia) have experienced a marked demographic decline starting in the 1990s and 2000s and still ongoing (around 5–7 generations, respectively; Quebec ministère des Forêts, de la Faune et des Parcs [MFFP], unpublished). RG herd increased from 61,842 individuals in 1963 (Des Meules & Brassard, 1964) to 823,000 ± 104,000 individuals in 1993 (Couturier, Courtous, Crépeau, Rivest, & Luttich, 1996). Then, it declined to 8,900 ± 668 individuals in 2016 (MFFP, unpublished), approximating a 99% decline in 23 years. For the RAF herd, it went from 56,000 individuals (Le Hénaff, 1976) to more than 628,000 individuals (1,193,000 ± 565,482; Couturier, Jean, Otto, & Rivard, 2004) between 1975 and 2001. It then decreased by about 70% to reach 199,000 (± 15,920) individuals in 2016 (MFFP; unpublished; Figure 1). Although the census sizes of RG and RAF herds are still relatively large compared to populations usually subjected to inbreeding depression (Blomqvist, Pauly, Larsson, & Flodin, 2010; Norén et al., 2016; Velando et al., 2015), certain mechanisms can reduce the effective size of these populations (i.e. the true number of effective breeders; Wright, 1931) and increase the risk of inbreeding. For instance, caribou is a polygynous species (L’Italien et al., 2012) that displays large variations of reproductive success generated by environmental (Couturier, Côté, Otto, Weladji, & Huot, 2009b) and individual heterogeneity (Pachkowski, Côté, & Festa-Bianchet, 2013), which can have such effect (Balloux et al., 2004; Wright, 1931). In addition, it has been shown that reindeer females do not avoid inbreeding when choosing a mating partner (Holand et al., 2007). Given the abrupt decline of RG and RAF herds and that caribou display characteristics that could increase inbreeding risk, it raises the question of whether inbreeding could have increased in the last decades and hence affected individual fitness in those populations, through inbreeding depression, and contributed further to their decline.

Here, we tested for a potential association between genomic inbreeding indices and two performance traits, body mass and annual
survival, in two declining herds of migratory caribou. Standardized multi-locus heterozygosity (sMLH) and $F_{grm}$ ($F^R$ in Yang et al., 2011) were estimated using 22,073 SNPs obtained by double-digest restriction-site-associated DNA sequencing (ddRADSeq). We hypothesized that inbreeding depression could result in positive correlations between genome-wide heterozygosity and performance traits (body mass and survival) and negative correlations between $F_{grm}$ and the same performance traits. Nevertheless, since those herds are still large (>8,000 individuals), we suspected that inbreeding depression could be low at this stage of demographic decline and have probably no lethal effects yet. Thus, higher correlations between genomic inbreeding indices and condition compared to survival were expected. Prior to analyses, we investigated the genetic differentiation between the two herds to confirm with SNP previous results obtained with microsatellites showing no differentiation between the herds (Boulet, Couturier, Côté, Otto, & Bernatchez, 2007).

2 METHODS

2.1 Study area and data collection

Using nets fired from helicopters, we captured and sampled 149 caribou of the RG herd (90 females; 59 males) between 2000 and 2014 and 251 caribou of the RAF herd (159 females; 92 males) between 1996 and 2016, following the guidelines from the Canadian Council on Animal Care. Individuals were captured up to five times over the study period, but most (85.5%) were captured only once. We collected muscle biopsies ($n = 361$), hairs ($n = 27$) or blood ($n = 12$) samples on each animal and froze all samples except for some biopsies ($n = 150$) that were stored in ethanol ($\geq$70%) at room temperature. We used EDTA to preserve blood samples and prevent them from coagulating. From the 400 individuals, 222 ($\geq 2$ years old) were fitted with tracking collars using different satellite networks (Argos, Iridium, Globalstar), allowing the assessment of annual survival from capture date to 2017. Mortalities were signalled by collars after 12–24 hr of inactivity and were then confirmed on the field based on visual cues when retrieving the collars. In addition, 253 caribou ($\geq 1$ year) were weighted to the nearest 0.1 kg using a hanging scale. Annual survival and body mass (collected in January, February, March, June, October or November) were considered as indices of individual performance (Taillon, Brodeur, Festa-Bianchet, & Côté, 2011) and were included in our models as response variables. Moreover, 55 calves (<2 days; Taillon, Barboza, & Côté, 2013) captured between 2007 and 2009 were used to ensure our sample (that did not include calves) was representative of heterozygosity levels at birth by comparing heterozygosity between calves and yearlings ($n = 49$) born in those years. Some of those calves ($n = 20$) were paired on the field with females included in our data set.

2.2 DNA extractions

We used Qiagen DNeasy Blood and Tissue Kit (Qiagen, Inc.) to extract DNA. To digest hairs (~100 hairs), we used 1.3 times the recommended volume of ATL lysis buffer and proteinase K and added 45 μl of dithiothreitol (DTT; 100 mg/ml). To maximize DNA yield, we eluted in 100 μl elution buffer for blood and hair samples and 200 μl otherwise; for all sample types, flow-through went through columns a second time at the final elution. We assessed DNA quality and checked for degradation on agarose gels. Most samples of blood or biopsies led to high-to-moderate DNA quality (low-to-medium.
degradation), whereas hair samples led to lower DNA quality with higher levels of degradation. We quantified DNA concentrations with a Qubit 2.0 Fluorometer (Life Technologies) and standardized all samples to 200 ng/μl before library construction.

### 2.3 ddRADSeq library construction and sequencing

We constructed ddRADSeq libraries from caribou’s genomic DNA following the general protocol from Peterson, Weber, Kay, Fisher, and Hoekstra (2012) with some modifications. Following recommendations of Mastretta-Yanes et al. (2015), we replicated 70 (17.5%) of our samples: 40 intra-library replicates and 30 inter-library replicates. We used a combination of 47 uniquely tagged P1 adaptors and 10 PCR indices to construct 10 libraries containing 47 DNA samples each. For each sample, 200 ng of DNA was digested with 20 units each of SbfI (CCTGCA/GG) and MspI (CC/GG). After ligation of P1 and P2 adaptors to SbfI and MspI cutsite overhang, respectively, we pooled all samples from one library and purified it using Agencourt AMPure XP (Beckman Coulter) or NucleoMag (Macherey-Nagel) systems with a DNA:magnetic bead solution ratio (v/v) of 1:1.8. We then selected fragments between 200 and 500 bp on agarose gels (1.6%), using Promega Wizard SV Gel and PCR Clean-Up System to perform gel extraction. Final amplification was achieved in 16 reactions of 20 μl containing 12.6 μl nanopure water, 4 μl Phusion© HF 5× Buffer (New England Biolabs), 0.4 μl dNTPs (10 mM), 0.4 μl each forward and reverse primers (10 μM; PCR1 and indexed PCR2), 0.2 μl Taq Phusion HF (2,000 U/ml) and 2 μl DNA template. After an initial denaturation step of 30 s at 98°C, 15 cycles of 10 s at 98°C, 10 s at 65°C and 30 s at 72°C were performed and followed by a final extension of 10 min at 72°C. All 16 PCRs from the same library were then pooled and purified using Promega Wizard SV Gel and PCR Clean-Up System. The 10 pooled libraries were each sequenced on two lanes of Illumina HiSeq 4000 (with HiSeq 3000/4000 SBS Kit) by Fasteris (Switzerland).

### 2.4 ddRADSeq data processing

We used Cutadapt (v. 1.8.1; Martin, 2011) to remove adapter sequences from sequence files and FastQC (v. 0.11.2; Andrews, 2010) to assess data quality and detect adaptor contamination. Then, we used Stacks (v. 1.44; Catchen, Amores, Hohenlohe, Cresko, & Postlethwait, 2011; Catchen, Hohenlohe, Bassham, Amores, & Cresko, 2013) to demultiplex data and build a de novo SNP catalog. We tested different sets of Stacks core parameters (ustack − m (2–6), − M (2–6) and − max_locus_stacks (2–6) and cstack − n (0–5)), by varying one parameter at a time whereas the others remained at their default values, as suggested by Mastretta-Yanes et al. (2015), and chose the values that minimized error rates between replicates (n = 70 pairs) and maximized the amount of data recovered. To do so, we combined files from the two sequencing runs to obtain a depth of coverage that would be similar to that observed in our final set, but

---

**FIGURE 2** Correlations between standardized multi-locus heterozygosity (sMLH) and individual harmonic mean of locus depth of coverage in relation to different values of upper bound after filtration for −lnl_lim = −10 in Rxstacks and without individuals that had more than 80% missing data. Stacks was run with optimal core parameter values identified by analysis of inter-replicates error rates (m = 3, M = 4, N = 6, max_locus_stacks = 3, n = 3). Free upper bound means that the model SNP was used in Ustacks instead of the bounded model.
kept only the replicate pairs at this stage. First, using the default SNP calling model, we identified the optimal values of \(-m\) (3), \(-M\) (4), \(-N\) (6), \(-\text{max}\_\text{locus}\_\text{stacks}\) (3) and \(-n\) (3; error rates for each set of parameters are presented in Figures S1 and S2). Then, to compare upper bound values for the bounded model to the default (SNP) model, we used those optimal values in combination with each value of upper bound tested (0.0125, 0.023, 0.5, 0.1, 0.15 and 1 [default SNP model]). In preliminary analysis, we found that genomic inbreeding indices were strongly correlated, in a nonlinear way, with the individual harmonic mean of locus depth of coverage. This trend was consistent in both herds and sex. We found that filtering data with \(-\text{lnl}\_\text{lim} = -10\) (95% of our loci had a log likelihood equal or higher than \(-10\)) in Rxstacks and eliminating individuals with more than 80% missing data made the relationship linear. We did so and then chose the value of upper bound that led to the weakest and least significant correlation between sMLH or \(F_{grm}\) (see ‘Assessment of genomic inbreeding indices’ for details) and individual harmonic mean of locus depth of coverage (Figures 2 and S3). The upper bound values of 0.0125 or 0.023 for the bounded model both led to very weak correlations that were nonsignificant for sMLH and barely significant for \(F_{grm}\) and we decided to use 0.023 as it also corresponded to the highest PhiX error rate reported for our sequencing runs. Note that error rates between replicates did not strongly vary when the upper bound value changed (result not shown). In all Stacks runs, including the one performed to produce our final set, we specified in the Populations component only one population, as RG and RAF herds were shown to be genetically undifferentiated (Boulet et al., 2007; result confirmed in the present study) and set \(-m\) (minimum stack depth required to call a genotype in an individual) to 8, \(-r\) (minimum percentage of individuals required to process a locus) to 0.5 and \(-\text{min}\_\text{maf}\) (minimum minor allele frequency required to process a nucleotide site at a locus) to 0.01. The \(-m\) value seemed to offer the best trade-off between the number of retained loci and genotype quality. We were rather permissive for \(-r\) and \(-\text{min}\_\text{maf}\) because the heterozygosity scores were standardized to account for missing data and mean observed heterozygosity at genotyped loci. This allowed retaining a maximum number of loci for subsequent analysis. To produce the final data set, we combined all replicate files in order to increase depth of coverage and genotype accuracy and ran Stacks with all parameters set to their optimal values identified above.

2.5 | Populations structure

To assess the genetic differentiation between the two herds, we filtered the data set described in ‘ddRADSeq data processing’ by keeping only a SNP by locus and using \(r = 0.8\) in Populations to avoid artificially increasing or decreasing differentiation between populations. This subset included 6,384 SNPs typed for 362 individuals (10.44% missing data). Prior to analysis, input files were exported in the appropriate format using radiator (v. 0.0.5; Gosselin, 2017). We removed from the data set known related individuals (e.g. mother-calf pairs identified in the field) or related individuals detected with a genetic relatedness matrix obtained using the ‘snpgdsGRM’ function implemented in the R SNPRelate package (Zheng et al., 2012). This resulted in a data set including 344 individuals. Genetic differentiation between herds was then investigated, using classical \(F\)-statistics according to Weir and Cockerham (1984) implemented in StAMPP (Pemberton, Cogan, & Forster, 2013), principal component analysis (PCA) and discriminant analysis of principal components (DAPC; Jombart, Devillard, & Balloux, 2010) implemented in adegenet (Jombart, 2008). DAPC was run on a priori defined clusters, that is, RG and RAF herd membership. We performed a cross-validation analysis to identify the optimal number of principal components (PCs) to retain in the DAPC analyses. To do so, we carried out a stratified cross-validation of DAPC using the function xvalDapc in adegenet, with 100 replicates at each level of PC retention and the data divided into two sets, that is, a training set (90% of the data) and a validation set (10% of the data). Afterwards, the optimal number of PCs was used in subsequent analysis.

2.6 | Assessment of genomic inbreeding indices

We computed individual sMLH with the function ‘sMLH’ implemented in the package inbredR (Stoffel et al., 2016) in R (v. 3.4.0; R Core Team, 2017). sMLH accounts for differences in the identity of the loci genotyped in all individuals and corrects the heterozygosity score for the mean heterozygosity observed at the typed loci in the rest of the population. Here, sMLH was computed for individuals from the two herds simultaneously. Then, using the ‘r2_hf’ function of the InbreedR package (Stoffel et al., 2016), we computed the expected correlation between sMLH and inbreeding \((F)\) with 100 bootstraps. We computed \(F_{grm}\) in GCTA (Yang et al., 2011), using the command ‘–ibc’. In addition, to ensure the reliability of our data sets to properly detect HFCs, we assessed identity disequilibrium (co-variance of heterozygosity among loci of a given individual; Weir & Cockerham, 1973; Szulkin et al., 2010) by computing the heterozygosity–heterozygosity correlation (HHC; Balloux et al., 2004) and the \(g_2\) value (David, Pujol, Viard, Castellas, & Goudet, 2007; Szulkin et al., 2010). HHC is a measure of correlation of heterozygosity across loci. The stronger the correlation is, the more the multi-locus heterozygosity is representative of global genetic diversity (Balloux et al., 2004). The \(g_2\) value measures the variance in the degree of covariance of heterozygosity at the individual level. The higher the variance is, the more likely it is to detect an HFC resulting from general effects in the sample if there is actually one in the population (David et al., 2007; Szulkin et al., 2010). HHC values were assessed using 100 iterations, and the \(g_2\) values were obtained using 100 permutations and 10 bootstraps, with the functions ‘HHC’ and ‘g2_snps’ of the InbreedR package (Stoffel et al., 2016). Because the data sets used to assess the effect of sMLH on body mass and annual survival did not include the same individuals, HHC and \(g_2\) were computed separately for both data sets. Although HHC and \(g_2\) are useful tools to detect identity disequilibrium, it was shown that HFCs could still occur and be caused by inbreeding depression in populations in which identity disequilibrium could not be detected (Kardos, Allendorf, & Luikart, 2014; Miller & Coltman, 2014).
2.7 Variation of genomic inbreeding indices through time and age

We suspected that the abrupt decline observed in the two herds could have caused a reduction of sMLH and an increase of $F_{\text{grm}}$ in more recent cohorts. We thus tested for variation of those genomic inbreeding indices through time by fitting linear models with either sMLH or $F_{\text{grm}}$ as response variables and cohort (birth year; coded as a continuous variable), herd and the interaction between the two as predictor variables. We used the argument mother–calf pairs (cut-off date 31 August 2008) were censored starting from the point of malfunction. We fitted generalized mixed-effects models with a binomial distributed error (logit link function), using the ‘glmer’ function implemented in the package lme4 (Bates et al., 2015). The baseline model included age, age$^2$, herd and sex (Couturier, Otto, Côté, Luther, & Mahoney, 2010; Loison, Festa-Bianchet, Gaillard, Jorgenson, & Jullien, 1999) as fixed-effect variables and year and individual identity as random factors. In the derived models, we added the simple (sMLH or $F_{\text{grm}}$) and quadratic (sMLH$^2$ or $F_{\text{grm}}^2$) terms of the genomic inbreeding indices and their interaction with sex, herd, age and age$^2$.

3 RESULTS

3.1 Population structure

We observed a significant but very low genetic differentiation between the RG and RAF herds ($F_{ST}$: 0.0027, 95% CI: [0.0024, 0.0029], p-value < .001). In line with this estimation, two genetic clusters appeared clearly on PCA (Figure 3a). Cross-validation determined that the use of the first 10 PCs for the DAPC was optimal. Although the 10 PCs portrayed only 12% of the variance, it gave the highest predictive success (88.9%) and the lowest root mean squared error (11.9%). Overall, DAPC indicated some genetic overlap between the two clusters, but 82.6% caribou of the RG herd and 92.1% caribou of the RAF herd were assigned to the correct herd (Figure 3b,c). The genetic relatedness matrix did not reveal strong particular individual clustering (results not shown), but confirmed that relatedness between known related individuals (i.e. mother–calf pairs of both herds; RG = 11 pairs and RAF = 5) was as expected ~0.5 (0.45 ± 0.04 [SD]).

3.2 Assessment of genomic inbreeding indices

Before the final SNP calling procedure, thirty-six of 400 individuals were excluded from the analyses because they presented more than 80% missing data. At the end of our de novo SNP calling procedure, we obtained 22,073 SNPs distributed at 14,847 loci. We then eliminated two individuals who had very high multi-locus heterozygosity scores before we computed sMLH as we suspected that they were contaminated during DNA extraction or library preparation. The median of individual locus depth of coverage was 22.08 reads/nonmissing locus. Individuals scored or library preparation. The median of individual locus depth of

2.8 Relationships between genomic inbreeding indices and performance

We used linear mixed-effects models with the function ‘lmer’ implemented in the R package lme4 (Bates, Mächler, Bolker, & Walker, 2015) to assess the relationship between genomic inbreeding indices and body mass. Body mass was log-transformed to improve homogeneity of variances, and sMLH and $F_{\text{grm}}$ were centred to improve convergence of models. To investigate the possible effect of genomic inbreeding indices in interaction with sex, age or herd, we adopted a model selection approach. We first considered a baseline model that included age, $\sqrt{\text{age}}$, sex, herd and month of capture as fixed-effect variables, as these factors are known to influence caribou body mass (Couturier, Côté, Huot, & Otto, 2009a; Couturier et al., 2009b; Parker, 1981). We added year and individual identity as random factors to consider annual variation and repeated measures on individuals. To the baseline model, we added sMLH and sMLH$^2$ or $F_{\text{grm}}$ and $F_{\text{grm}}^2$ as fixed-effect variables to build simple models that included genomic inbreeding indices. Considering a potential quadratic effect of the genomic inbreeding indices on performance, we accounted for a potential threshold under which a decrease in inbreeding would not have any effect on performance traits. Then, we derived competitive models including interactions between the simple (sMLH or $F_{\text{grm}}$) and quadratic (sMLH$^2$ or $F_{\text{grm}}^2$) terms of genomic inbreeding indices and sex, herd, age and $\sqrt{\text{age}}$. All models were compared using Akaike’s information criterion for small sample sizes (AICc), and we selected the model with the lowest value of AICc. $\Delta$AICc and AICc weights ($\omega_0$) were computed with the $R$ package AICmodavg (Mazerolle, 2017). We used the same approach to test for the effect of genomic inbreeding indices on annual survival in adults (>2 years). Annual survival was coded as a binary variable with a value of 1 if the individual was alive during a given year and 0 if the individual was dead. Malfunctioning collars were censored starting from the point of malfunction. We fitted generalized mixed-effects models with a binomial distributed error (logit link function), using the ‘glmer’ function implemented in the package lme4 (Bates et al., 2015). The baseline model included age, age$^2$, herd and sex (Couturier, Otto, Côté, Luther, & Mahoney, 2010; Loison, Festa-Bianchet, Gaillard, Jorgenson, & Jullien, 1999) as fixed-effect variables and year and individual identity as random factors. In the derived models, we added the simple (sMLH or $F_{\text{grm}}$) and quadratic (sMLH$^2$ or $F_{\text{grm}}^2$) terms of the genomic inbreeding indices and their interaction with sex, herd, age and age$^2$. We observed a significant but very low genetic differentiation between the RG and RAF herds ($F_{ST}$: 0.0027, 95% CI: [0.0024, 0.0029], p-value < .001). In line with this estimation, two genetic clusters appeared clearly on PCA (Figure 3a). Cross-validation determined that the use of the first 10 PCs for the DAPC was optimal. Although the 10 PCs portrayed only 12% of the variance, it gave the highest predictive success (88.9%) and the lowest root mean squared error (11.9%). Overall, DAPC indicated some genetic overlap between the two clusters, but 82.6% caribou of the RG herd and 92.1% caribou of the RAF herd were assigned to the correct herd (Figure 3b,c). The genetic relatedness matrix did not reveal strong particular individual clustering (results not shown), but confirmed that relatedness between known related individuals (i.e. mother–calf pairs of both herds; RG = 11 pairs and RAF = 5) was as expected ~0.5 (0.45 ± 0.04 [SD]).
confirming the power of our marker sets to detect HFCs. $F_{grm}$ ranged between $-0.16$ and $0.05$. We did not detect differences of sMLH between herds (RAF herd mean ± SE: $1.001 ± 0.001$ and RG herd mean ± SE: $0.999 ± 0.002$; two-sample $t$ test $t = 1.170$, $df = 432$, $p$-value $= .24$), nor significant differences of $F_{grm}$ (RAF herd mean ± SE: $-0.026 ± 0.001$ and RG herd mean ± SE: $-0.025 ± 0.002$; Wilcoxon rank-sum test $W = 21,015$, $p$-value $= .58$). The average (unstandardized) observed heterozygosity was $0.23 ± 0.01$ (SD) in both herds and ranged between $0.20$ and $0.27$. The HHC values were $0.30 ± 0.05$ (SD) for the mass data set and $0.33 ± 0.05$ (SD) for the survival data set. The $g_z$ values were $3.8 \times 10^{-4} ± 4.4 \times 10^{-4}$ (SE; $p(g_z > 0) = 1$) for the mass data set and $1.2 \times 10^{-4} ± 3.8 \times 10^{-4}$ (SE; $p(g_z > 0) = 1$) for the survival data set.

### 3.3 Variation of genomic inbreeding indices through time and age

sMLH and $F_{grm}$ did not vary significantly through the sampling period (Figures S5 and S6) or between herds (Tables S1 and S2). We found no significant difference in sMLH or $F_{grm}$ between individuals born between 2007 and 2009 and captured for the first time either as calves or as yearlings (Tables S3 and S4).

### 3.4 Relationships between genomic inbreeding indices and performance

Model selection performed with sMLH and $F_{grm}$ produced similar results. In the models, $F_{grm}$ had estimates of the opposite sign...
compared to sMLH, as the two variables are negatively correlated, but significance was similar. Only results for sMLH candidate sets and models are presented here, and results for \( F_{\text{grm}} \) are detailed in Electronic Supplementary Material (Tables S5 and S6). Regarding sMLH model selection, the baseline model for the body mass candidate set was the best model (\( \Delta \text{AICc} = 0 \) and \( \omega_i = 0.49; \Delta \text{AICc} \) of the second-ranked model = 3.1; Table S7). Age, \( \sqrt{\text{age}} \), herd, sex and some months of capture had a significant effect on body mass in the baseline model (Table 1). Body mass increased from 1 to 2 years and plateaued after that age at around 90 kg for females and 106 kg for males. Individuals from the RG herd were on average 8.07 kg heavier than their RAF counterparts in June, and all individuals were on average heavier in fall and lighter at the beginning of summer, compared to winter months. Next candidate models did not include significant effects of sMLH or sMLH\(^2\) and their interactions (Figure S7), nor of \( F_{\text{grm}} \) or \( F_{\text{grm}}^2 \) and their interactions.

The baseline model was also selected as the best model for survival (\( \Delta \text{AICc} = 0 \) and \( \omega_i = 0.55; \Delta \text{AICc} \) of the second-ranked model = 3.3; Table S8). However, only herd had a significant effect on survival (Table 2), which was lower for individuals from the RG herd in comparison with RAF herd (Odds ratio = 0.46, 95% CI: [0.27, 0.78]; Figure 5). None of the next-ranked models included significant effect of sMLH or its interactions.

**DISCUSSION**

### 4.1 Population structure

Contrarily to what has been shown previously (Boulet et al., 2007; Yannic et al., 2016), we found a significant but very low differentiation between RG and RAF herds. The large number of SNPs used in our study improved population structure delineation in this context of weak genetic structure (see also Benestan et al.,...
and likely improved the detection of differences compared to microsatellite markers (Fischer et al., 2017; Gärke et al., 2011; Liu, Chen, Wang, Oh, & Zhao, 2005). In addition, it appears that the recent demographic decline of both populations has been accompanied by a reduction in their overlapping ranges (Le Corre, 2016; Figure 1) that could have contributed to reduce gene flow.

Nevertheless, the overall very low level of differentiation we observed and the few cases of herd switching observed in our data set are concordant with previous observations of high gene flow between RG and RAF herds (Boulet et al., 2007; Yannic et al., 2016).

### 4.2 Assessment of genomic inbreeding indices

It was suggested that SNP markers would greatly improve HFC studies by increasing representativeness of genome-wide diversity (Kardos et al., 2016; Miller et al., 2014). Accordingly, it was shown in harbour seal (Phoca vitulina) that the strength of an HFC detected in a natural population increased significantly when heterozygosity was estimated with 14,585 SNP markers rather than 27 microsatellites (Hoffman et al., 2014). Although we found higher values of HHC in our data sets than those usually reported in HFC studies using microsatellites (Guinand et al., 2013; Queirós, Vicente, Alves, de la Fuente, & Gortazar, 2016; Voegeli, Saladin, Wegmann, & Richner, 2013), our values were somewhat small, indicating that our estimate of heterozygosity may not have been representative of inbreeding levels (Balloux et al., 2004). In preliminary analyses, we found that some of the filters we applied in Stacks to reconstruct SNPs helped reduce the unintended correlation between genomic inbreeding indices and depth of coverage, but also reduced dramatically the values of HHC. HHC values should be high if markers’ heterozygosity was highly dependent on a third variable, like the depth of coverage. Thus, it was not surprising that HHC values drastically decreased as the relationship between sMLH and depth of coverage weakened.

The effect of the number of markers used to assess heterozygosity (Miller & Coltman, 2014) or marker type and population history (Miller et al., 2014) on identity disequilibrium has been studied previously, but the effect of depth of coverage has not. Here, we used error rates between replicate samples (Mastretta-Yanes et al., 2015) to calibrate filtering parameters in Stacks, but those were not accurate indicators of the correlation between depth of coverage and genomic inbreeding indices. Further studies should explore the impact of bioinformatic filters on SNP-based estimates of multi-locus heterozygosity, \( F_{grm} \) and identity disequilibrium and, in the context of HFC studies, determine the best approach to adopt for data filtering. Meanwhile, SNPs could still improve HFC studies because they provide a broader representation of general genetic diversity than microsatellites by covering coding and noncoding regions of the genome (Balloux et al., 2004; Miller & Coltman, 2014; Miller et al., 2014).

### 4.3 Relationships between genomic inbreeding indices and performance

We used HFCs to detect inbreeding depression in RG and RAF declining herds of migratory caribou. We studied the association of

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>SE</th>
<th>Z-value</th>
<th>p-value</th>
<th>95% Confidence interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.44</td>
<td>0.33</td>
<td>7.5</td>
<td>.00</td>
<td>(1.80, 3.07)</td>
</tr>
<tr>
<td>Centred age</td>
<td>−0.25</td>
<td>0.16</td>
<td>−1.6</td>
<td>.10</td>
<td>(−0.56, 0.05)</td>
</tr>
<tr>
<td>(Centred age)(^2)</td>
<td>−0.07</td>
<td>0.07</td>
<td>−0.9</td>
<td>.36</td>
<td>(−0.21, 0.08)</td>
</tr>
<tr>
<td>Sex—male</td>
<td>−0.17</td>
<td>0.26</td>
<td>−0.7</td>
<td>.51</td>
<td>(−0.69, 0.34)</td>
</tr>
<tr>
<td>Herd—Rivière-George</td>
<td>−0.77</td>
<td>0.27</td>
<td>−2.8</td>
<td>.01</td>
<td>(−1.30, −0.24)</td>
</tr>
</tbody>
</table>

Lower bounds and upper bounds of the 95% confidence intervals on estimates are provided in addition to the standard errors (SE), Z-values and p-values. Variables with 95% CI that do not overlap 0 are considered to have a significant effect and are in bold. Levels of reference are Female for sex and Rivière-aux-Feuilles for herd. Baseline model: annual survival ~ age + age \(^2\) + sex + herd + (1|year) + (1|individual identity).
genome-wide, multi-locus heterozygosity and $F_{grm}$, an inbreeding coefficient based on genomic relatedness, inferred with 22,730 SNPs with two performance traits: body mass and annual survival. The effects of nongenetic variables on body mass and survival conform to what had been found previously (Couturier et al., 2009a,b; Parker, 1981), and we did not find any effect of heterozygosity, nor inbreeding coefficient, on these traits. On the contrary, positive effects of heterozygosity on reproductive success and survival have been found in a population of European shag (Phalacrocorax aristotelis) that presented similar levels of demographic decline as the RAF herd (Velando et al., 2015), but also high philopatry and a much smaller population size (a few hundreds). Those factors have probably increased the probability of inbreeding in this population (Velando et al., 2015). In our study, genomic inbreeding indices remained stable over time. It is likely that the genetic diversity of the herds was shaped by the low population sizes preceding their augmentation in the 1970s and 1980s. Thus, it is expected that the return to similar sizes would not lead to decrease in population genetic diversity (e.g. in terms of number of alleles). Regarding individual genetic diversity, it was suggested that relaxed polygyny in caribou could attenuate the negative effect of population decline on genetic diversity as it increases the number of breeders in the population (Lovatt & Hoelzel, 2014). The relatively large population sizes (a few to many thousands of individuals) also contribute largely to reduce inbreeding risk.

We are confident that our marker set reflected reliable biological information, as we found estimations of relatedness between known related individuals (i.e. mother–calf pairs of both herds) were as expected ~0.5. Nevertheless, the $g_2$ values were not significantly different from 0, suggesting that the variance in inbreeding in our sample was too small to detect an HFC. This result could indicate that our sample did not capture all the variance observed in the herds, but we made our sample as representative as it could, a priori, by sampling 400 individuals of both sexes and various ages from the two herds, from different cohorts and in different seasons of a 20-year period. In addition, we checked that our sample was representative of the populations before potential selection against inbred individuals could occur and eliminate variance in heterozygosity. We found no significant difference in sMLH scores or inbreeding coefficient between calves and yearlings born the same years, suggesting that our sample was representative of the population, even though we did not include calves in our models to assess the effect of genomic inbreeding indices. Thus, we believe that the low variance in inbreeding in our sample reflected the true levels of variance in the populations. In this study, the average (unstandardized) observed heterozygosity was $0.23 \pm 0.01$ (SD) for the RG herd and $0.23 \pm 0.01$ (SD) for the RAF herd. In another study conducted on those populations, observed heterozygosity estimated at 7 microsatellite markers was 0.71 for the RG herd ($n = 98$) and 0.73 for the RAF herd ($n = 114$, Boulet et al., 2007). Similar results were more recently obtained with a larger panel of 16 microsatellite markers: 0.76 for the RG herd ($n = 71$) and 0.75 for the RAF herd ($n = 77$; G. Yannic, unpublished).

Since microsatellites are usually much more polymorphic than SNPs, because of their high mutation rate, their high number of alleles and the large ascertainment bias in polymorphic rates (Queiros et al., 2014), we expected to observe lower heterozygosity estimates with SNPs (Kaiser et al., 2017). Here, the values of heterozygosity for the two types of markers are comparable to estimates of heterozygosity made in a population of bighorn sheep (Ovis canadensis; $n = 26$) after a successful genetic rescue, where observed heterozygosity was 0.28 when estimated with 412 SNP markers and 0.64 when estimated with 200 microsatellite markers (Hogg, Forbes, Steele, & Luikart, 2006; Miller et al., 2014). Thus, assuming that our sample was representative of studied populations, our results suggest that RG and RAF herds do not suffer from inbreeding. Unless there were substructure and nonrandom mating within RG and RAF herds, which is not supported by our data, or high genetic loads, which we did not address here, it is expected that relatively large populations with a few to many thousand individuals would not suffer from inbreeding depression.

### 5 | CONCLUSION

In this study, we tested for the relationship between sMLH or $F_{grm}$, both proxies of inbreeding, and body mass and survival in caribou. Using a large number of markers (~22k SNPs), we did not detect an association of the genomic inbreeding indices with body mass or annual survival. Furthermore, sMLH and $F_{grm}$ remained stable over the period monitored, which suggests that the rapid and intense demographic decline of the herds did not cause inbreeding depression in those populations. Additionally, we found a strong bias in individual heterozygosity associated with depth of sequencing coverage, which, once corrected for, reduced identity disequilibrium in our data set and may have reduced our ability to detect HFCs. This result highlights that depth of coverage should be taken into consideration when assessing heterozygosity and that, more generally, some questions remain regarding suitable filtration of SNP data, especially in HFC studies.

Although we found no evidence for HFCs, the strong rate of decline observed in RG and RAF herds could have and could continue to increase levels of genetic drift and lead to the loss of adaptive genetic variation (Allendorf, 1986; Taylor, Jenkins, & Arcese, 2012), or to the accumulation of deleterious mutations (Lohr & Haag, 2015; Perrier, Ferchaud, Sirois, Thibault, & Bernatchez, 2017) at the population level. Further studies should focus on direct HFCs in those populations to understand how such processes could impair individual performance and limit evolutionary potential, especially in the face of climate change, where genetic diversity would provide a selective advantage (Allendorf et al., 2013; Forcada & Hoffman, 2014).

### ACKNOWLEDGMENTS

We thank C. Miquel, D. Rioux, L. Denoyelle, T. Capblanc, N. Curt-Grand-Gaudin and N. Tissot as well as the AnaBM (USMB) and AEEM (UGA) laboratory facilities for their help and support during
the laboratory work. We acknowledge C. Landry, J. Turgeon, T. Bilde, deciding editor, and two anonymous reviewers for their comments that helped improve the quality of this manuscript and M. LeCorre for his help with Figure 1. We are indebted to the ministère des Forêts, de la Faune et des Parcs (MFFP) of Québec for capture and monitoring of caribou, especially S. Rivard, V. Brodeur and J. Taillon. We are grateful for the logistic and financial resources made available by Caribou Ungava and the MFFP. The funding of Caribou Ungava is provided by the Natural Sciences and Engineering Research Council of Canada, the MFFP, ArcticNet, the Department of Environment and Conservation of Newfoundland and Labrador, Tata Steel Minerals Canada Limited, Hydro-Québec, GlenCore-Mine Raglan, Grand Council of the Crees, Tornagat Wildlife, Plants and Fisheries Secretariat, Société Makivik—Air Inuit, Exploration Osisko Baie-James, Azimut Exploration and Fédération des Pourvoiries du Québec. Funding for laboratory work was also provided by a LabEx OSUG@2020 grant. M. Gagnon was supported by the Natural Sciences and Engineering Research Council of Canada, the Fonds de Recherche Nature et Technologies du Québec and Caribou Ungava.

DATA AVAILABILITY

Genetic data and some ecological data are available at https://doi.org/10.5061/dryad.crdjfn30k. Additional data could be made available on demand.

ORCID

Marianne Gagnon https://orcid.org/0000-0001-8125-312X
Glenn Yannic https://orcid.org/0000-0002-6477-2312
Steve D. Côté https://orcid.org/0000-0002-4875-1917

REFERENCES


**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section at the end of the article.

**How to cite this article:** Gagnon M, Yannic G, Perrier C, Côté SD. No evidence of inbreeding depression in fast declining herds of migratory caribou. *J Evol Biol*. 2019;00:1–14. https://doi.org/10.1111/jeb.13533